V Namratha
1NT19IS180

Program - 4:

```
package nam;

import java.io.IOException;
import java.util.Iterator;

import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapred.FileInputFormat;
import org.apache.hadoop.mapred.FileOutputFormat;
import org.apache.hadoop.mapred.JobClient;
import org.apache.hadoop.mapred.JobConf;
import org.apache.hadoop.mapred.MapReduceBase;
import org.apache.hadoop.mapred.Mapper;
import org.apache.hadoop.mapred.OutputCollector;
import org.apache.hadoop.mapred.Reducer;
import org.apache.hadoop.mapred.Reporter;
import org.apache.hadoop.mapred.TextInputFormat;
import org.apache.hadoop.mapred.TextOutputFormat;

public class TransactionCount {

    public static class Map extends MapReduceBase implements
    Mapper<LongWritable, Text, Text, IntWritable> {
    private final static IntWritable one = new IntWritable(1);
    //private Text word = new Text();

    public void map(LongWritable key, Text value, OutputCollector<Text,
    IntWritable> output, Reporter reporter) throws IOException {
    String myString = value.toString();
    String[] userCount = myString.split(",");
    output.collect(new Text(userCount[3]), one);

    }
    }
```
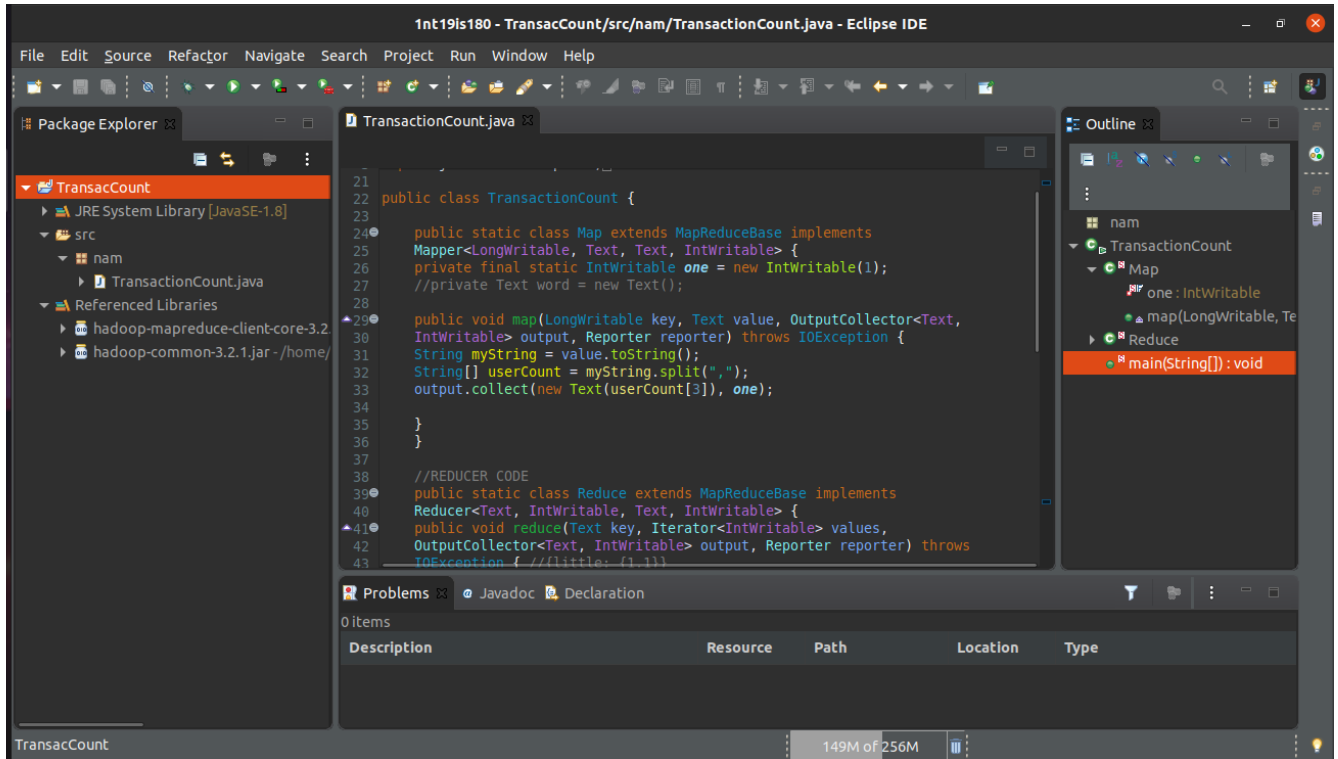
```java
//REDUCER CODE
public static class Reduce extends MapReduceBase implements
Reducer<Text, IntWritable, Text, IntWritable> {
public void reduce(Text key, Iterator<IntWritable> values,
OutputCollector<Text, IntWritable> output, Reporter reporter) throws
IOException { //{little: {1,1}}
int finaluserCount = 0 ;
Text mykey = key ;
while(values.hasNext()) {
IntWritable value = values.next();
finaluserCount += value.get();
}
output.collect(mykey, new IntWritable(finaluserCount));
}
}
//DRIVER CODE
public static void main(String[] args) throws Exception {
JobConf conf = new JobConf(TransactionCount.class);
conf.setJobName("wordcount");
conf.setOutputKeyClass(Text.class);
conf.setOutputValueClass(IntWritable.class);
conf.setMapperClass(Map.class);
conf.setCombinerClass(Reduce.class);
conf.setReducerClass(Reduce.class);
conf.setInputFormat(TextInputFormat.class);
conf.setOutputFormat(TextOutputFormat.class);
FileInputFormat.setInputPaths(conf, new Path(args[0]));
FileOutputFormat.setOutputPath(conf, new Path(args[1]));
JobClient.runJob(conf);
}

}
```
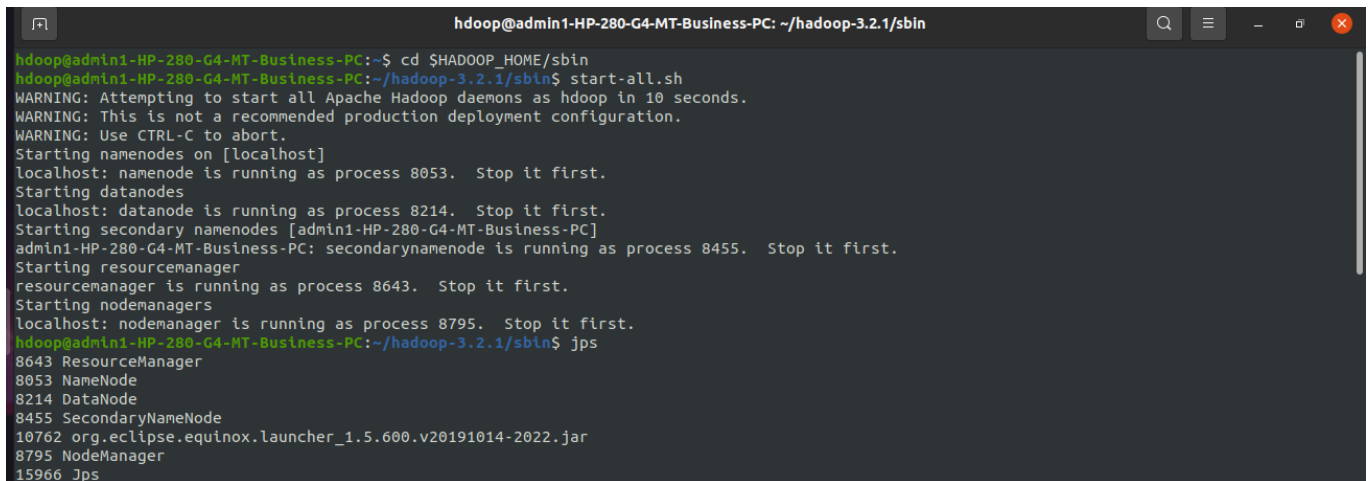
```
cd $HADOOP_HOME/sbin
start-all.sh
```



```
jps
hdfs dfs -mkdir -p /kpoppppp
hdfs dfs -copyFromLocal /home/hdoop/Documents/nams.csv /kpoppppp
hdfs dfs -cat /zpopppppp/part*
```

```
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ jps
8643 ResourceManager
8053 NameNode
8214 DataNode
8455 SecondaryNameNode
10762 org.eclipse.equinox.launcher_1.5.600.v20191014-2022.jar
8795 NodeManager
15966 Jps
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hdfs dfs -mkdir -p /kpoppppp
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hdfs dfs -copyFromLocal /home/hdoop/Documents/nams.csv /kpoppppp
2022-06-23 15:03:34,175 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hadoop jar /home/hdoop/Desktop/namru19999.jar /kpoppppp /zpoppppp
2022-06-23 15:04:31,649 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2022-06-23 15:04:31,772 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2022-06-23 15:04:31,906 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and
 execute your application with ToolRunner to remedy this.
2022-06-23 15:04:31,933 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hdoop/.staging/job_165
5973783279_0001
2022-06-23 15:04:32,021 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-23 15:04:32,122 INFO mapred.FileInputFormat: Total input files to process : 1
2022-06-23 15:04:32,165 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-23 15:04:32,282 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-23 15:04:32,290 INFO mapreduce.JobSubmitter: number of splits:2
```

hdfs dfs -cat /zpoppppp/part*

```
                Input split bytes=182
                Combine input records=6
                Combine output records=5
                Reduce input groups=4
                Reduce shuffle bytes=73
                Reduce input records=5
                Reduce output records=4
                Spilled Records=10
                Shuffled Maps =2
                Failed Shuffles=0
                Merged Map outputs=2
                GC time elapsed (ms)=155
                CPU time spent (ms)=1150
                Physical memory (bytes) snapshot=784547840
                Virtual memory (bytes) snapshot=7618490368
                Total committed heap usage (bytes)=828375040
                Peak Map Physical memory (bytes)=297181184
                Peak Map Virtual memory (bytes)=2538024960
                Peak Reduce Physical memory (bytes)=191840256
                Peak Reduce Virtual memory (bytes)=2544308224
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
                IO_ERROR=0
                WRONG_LENGTH=0
                WRONG_MAP=0
                WRONG_REDUCE=0
        File Input Format Counters
                Bytes Read=203
        File Output Format Counters
                Bytes Written=33
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hdfs dfs -cat /zpoppppp/part*
2022-06-23 15:05:21,006 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
10000   1
20000   3
30000   1
Amount  1
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$
```